!!! Manuel en construction !!!
Crochets [...] mis pour « à compléter »

ANALEC : MANUEL D'UTILISATION

Table des matières

I. Avant de commencer

1. Télécharger Analec

Aller sur le site du labo : http://www.lattice.cnrs.fr

L'exécutable ses trouve sous : **Productions scientifiques> Logiciels > Analec**

Vous y accéder directement à cette adresse : http://www.lattice.cnrs.fr/analec

Actuellement la version disponible est la : 1.4

Télécharger **Executable version 1.4**

Et placer le fichier sur son ordi dans le dossier de travail souhaité : Exemple :

C:\Users\Frederique\LaTTiCe\Analec\

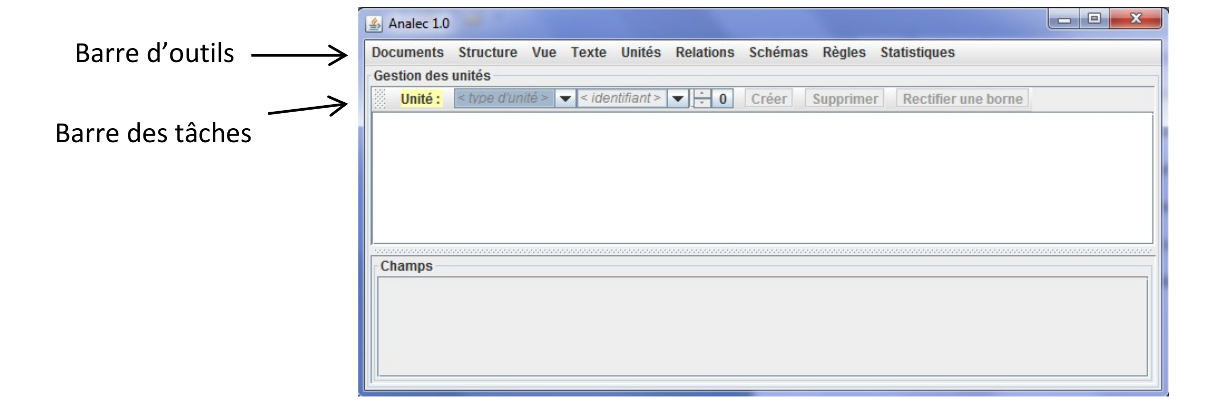
Dans le dossier de travail se retrouve alors le fichier Analec-6.jar[1]

Il s'agit de l'exécutable qui vous permettra de lancer l'outil d'annotation Analec.

2. Ouvrir Analec

Ouvrir Analec, en cliquant sur : **Analec-6.jar**

[1] Le fichier Analec-6.jar ouvre la version 1.4 du logiciel – erreur du nom de fichier, à modifier.



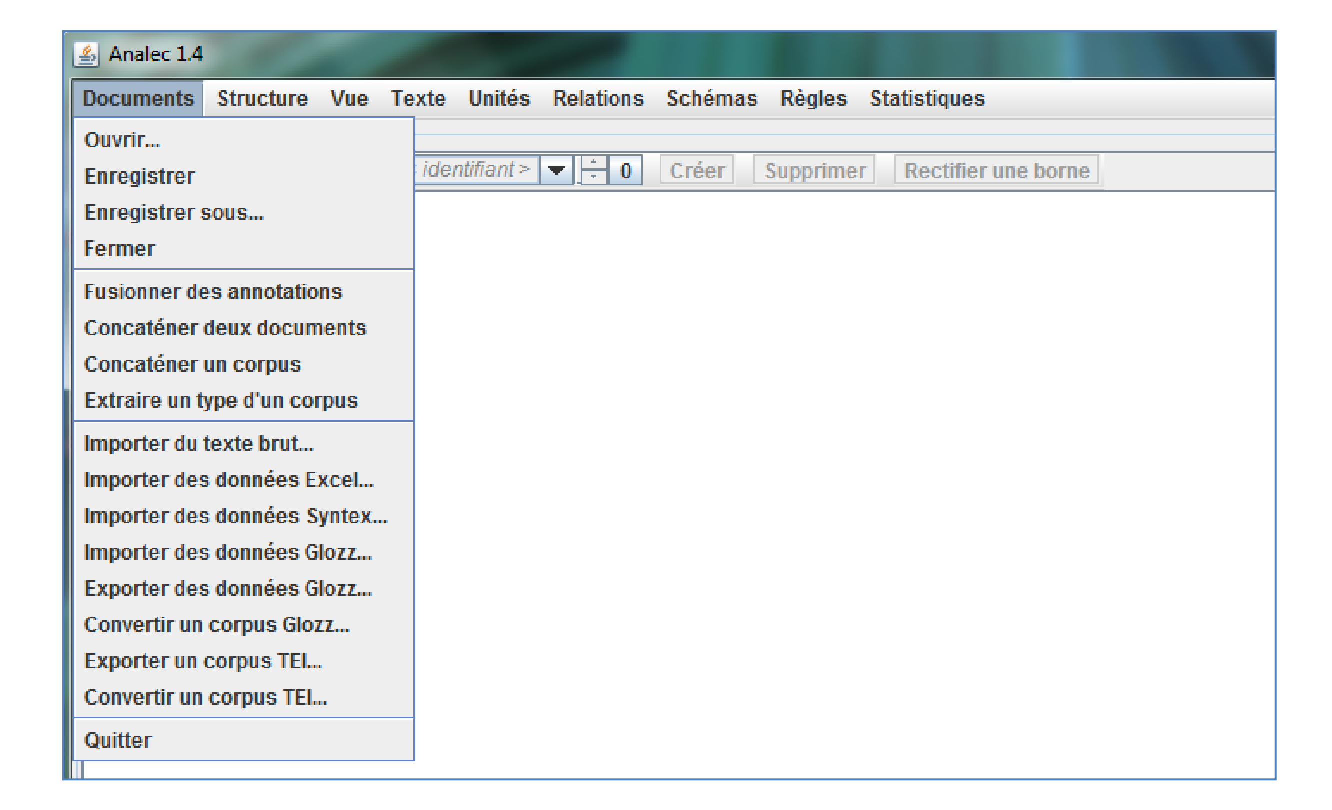
! Pour exécuter ce fichier, la seule nécessité est d'avoir un environnement java disponible sur son ordinateur (.jrc).

II. Les fonctionnalités d'Analec

Nous prenons ci-dessous successivement chacun des titres de la barre d'outils.

1. La barre d'outils

a) Documents



Le menu est divisé en quatre zones. Chacune d'elles permet d'agir sur un certain type de fichiers.

- **… à partir d'un document Analec.**

zone située en haut du menu.

Ouvrir un document Analec.
Les documents Analec ont une extension **.ec**.
Ouvrir un fichier qui a déjà était ouvert et enregistré sous Analec.

Enregistrer un document de travail.
! Le raccourci CTL-s existe.
! Une étoile apparaît à la suite du nom du fichier quand les dernières modifications apportées ne sont pas enregistrées.

Enregistrer sous Faire une copie du document courant sous le nom souhaité.
! comme dans de nombreux logiciels, c'est la copie qui reste ouverte et sur laquelle l'utilisateur va continuer à travailler.

- **… à partir de plusieurs documents Analec.**

Zone centrale du menu.
[…]

Concaténer un corpus Fonction qui concatène/regroupe en un seul fichier un ensemble de fichiers (l'ensemble des fichiers *.ec du répertoire sélectionner).

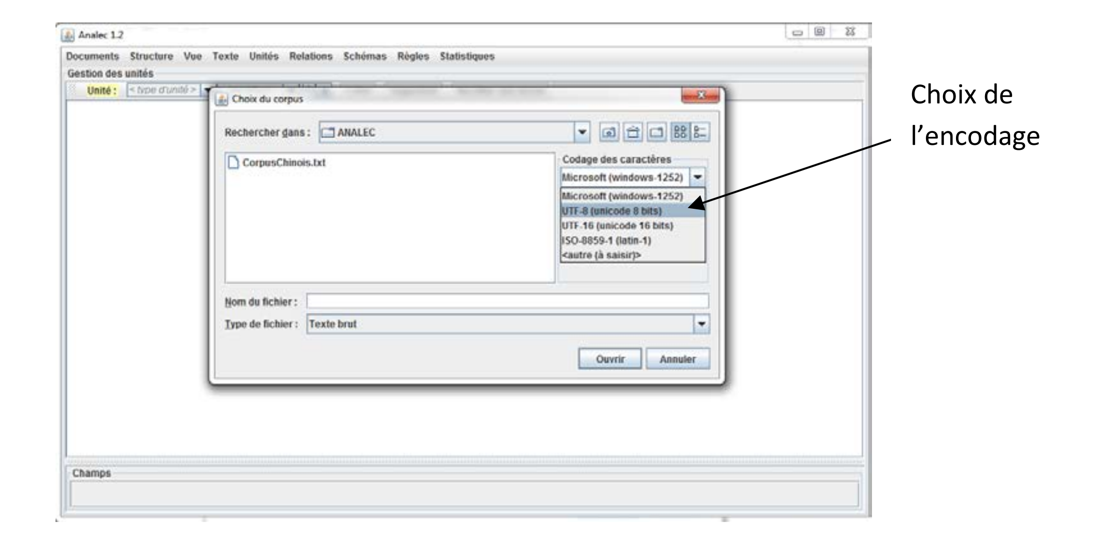
! Suite à cette action, c'est le fichier concaténé qui reste ouvert dans l'outil.
! Fonction pratique pour avoir des statistiques globales sur un ensemble de fichiers.
Attention : l'objet traité n'est plus le fichier dans son individualité, mais bien le corpus, c'est-à-dire l'ensemble des fichiers.

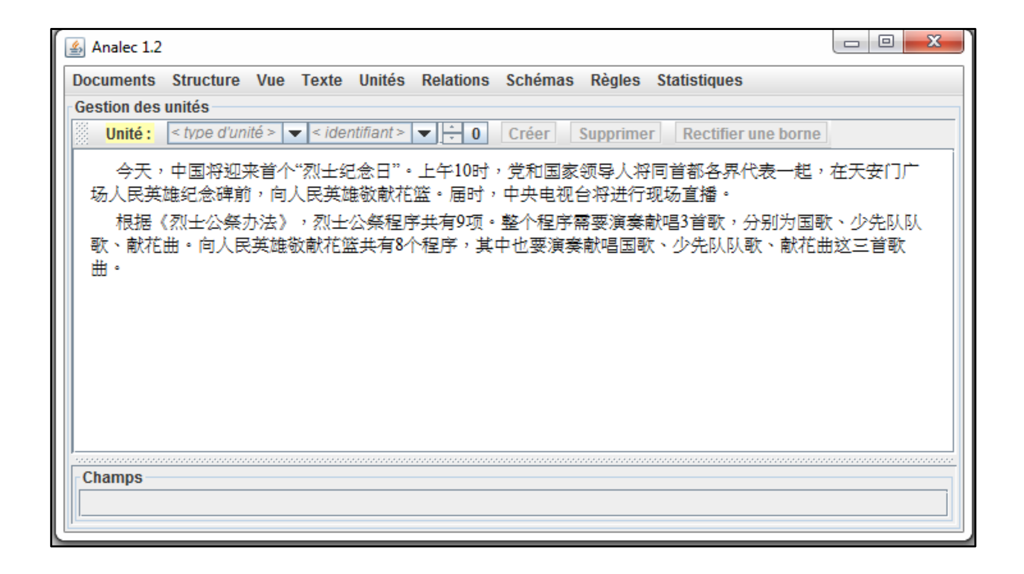
- **… à partir de documents divers - autres qu'au format Analec.**

Zone située en bas du menu.

importer un texte brut permet d'importer un fichier .txt

! Il est possible d'indiquer à Analec le type d'encodage du texte à importer.





[…]

Importer des données Glozz

Si des textes ont été annotés dans Glozz, il est possible ici de les importer dans Analec.

Une boîte de dialogue s'ouvre, il sera demandé successivement à l'utilisateur d'indiquer :

- le fichier texte à importer, c'est-à-dire le **c**orpus : *.a**c**
- le fichier contenant les **a**nnotations : *.a**a**
- le fichier contenant le modèle d'annotation : *.aa**m**

Attention :

Le fichier *.aa contient uniquement les annotations présentes dans le corpus annoté, celles qui ont été utilisées. Pour avoir la totalité du schéma d'annotation – et donc aussi les

étiquettes non encore utilisées dans le corpus annoté, il faut importer le modèle Glozz :

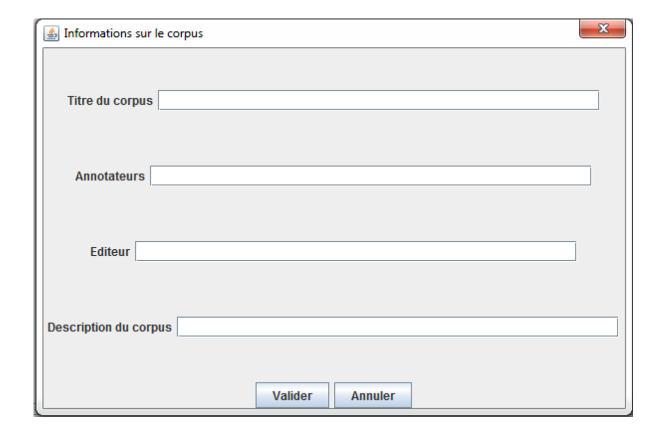
structure> Importer un modèle Glozz > *.aam

Utilité du fichier *modèle* :
Ce fichier permet à un groupe travaillant sur un même projet, une même problématique d'utiliser le même modèle d'annotation.

[...]

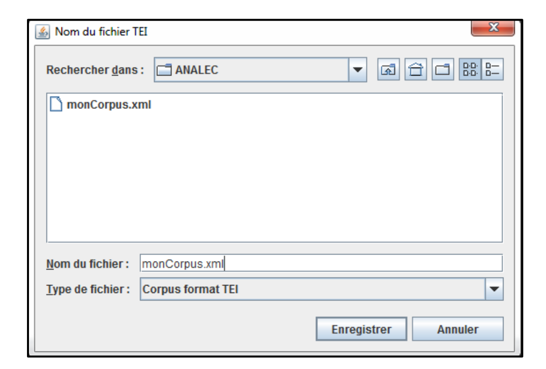
Exporter un corpus TEI…

Un texte annoté sous Analec (*.ec) peut être exporté au format XML. Le format est produit automatiquement par l'outil à partir des annotations réalisées. Pour générer ce format, l'utilisateur est invité à donner :
- un titre au corpus
- le nom du ou des annotateur(s)
- un éditeur
- une brève description du corpus



! Le remplissage de chacun de ces champs est facultatif.

Dans une seconde fenêtre, il doit indiquer le titre du document (info obligatoire) :



Le format XML est utile pour une exploitation automatisée des annotations et permet une conservation pérenne des données.

```
<p>
Lorsque la guerre fut déclarée, <anchor xml:id="u-Maillon-61-start"
type="AnalecDelimiter" subtype="UnitStart"/>le fils Sauvage, <anchor xml:id="u-
Maillon-62-start" type="AnalecDelimiter" subtype="UnitStart"/>qui<anchor xml:id="u-
Maillon-62-end" type="AnalecDelimiter" subtype="UnitEnd"/> avait alors trente-trois
ans<anchor xml:id="u-Maillon-61-end" type="AnalecDelimiter" subtype="UnitEnd"/>,
<anchor xml:id="u-Maillon-63-start" type="AnalecDelimiter"
subtype="UnitStart"/>s&apos;<anchor xml:id="u-Maillon-63-end"
type="AnalecDelimiter" subtype="UnitEnd"/>engagea, <anchor xml:id="u-Maillon-64-
start" type="AnalecDelimiter" subtype="UnitStart"/>laissant<anchor xml:id="u-
Maillon-64-end" type="AnalecDelimiter" subtype="UnitEnd"/> <anchor xml:id="u-
Maillon-65-start" type="AnalecDelimiter" subtype="UnitStart"/>la mère<anchor
xml:id="u-Maillon-65-end" type="AnalecDelimiter" subtype="UnitEnd"/> seule au
logis. <anchor xml:id="u-Maillon-66-start" type="AnalecDelimiter"
subtype="UnitStart"/>On<anchor xml:id="u-Maillon-66-end" type="AnalecDelimiter"
subtype="UnitEnd"/> ne <anchor xml:id="u-Maillon-67-start" type="AnalecDelimiter"
subtype="UnitStart"/>la<anchor xml:id="u-Maillon-67-end" type="AnalecDelimiter"
subtype="UnitEnd"/> plaignait pas trop, <anchor xml:id="u-Maillon-68-start"
type="AnalecDelimiter" subtype="UnitStart"/>la vieille<anchor xml:id="u-Maillon-68-
end" type="AnalecDelimiter" subtype="UnitEnd"/>, parce qu&apos;<anchor xml:id="u-
Maillon-69-start" type="AnalecDelimiter" subtype="UnitStart"/>elle<anchor
xml:id="u-Maillon-69-end" type="AnalecDelimiter" subtype="UnitEnd"/> avait de
l&apos;argent, <anchor xml:id="u-Maillon-70-start" type="AnalecDelimiter"
subtype="UnitStart"/>on<anchor xml:id="u-Maillon-70-end" type="AnalecDelimiter"
subtype="UnitEnd"/> le savait.
</p>
```

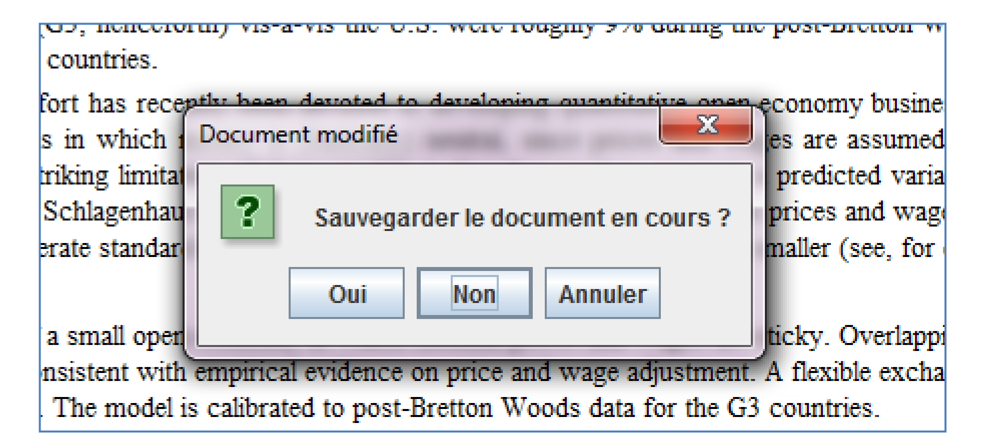
[…]

- **Quitter**

Cette action permet de fermer Analec.

A noter :

Si des modifications apportées au document ne sont pas enregistrées, un message d'avertissement est généré.



b) Structure

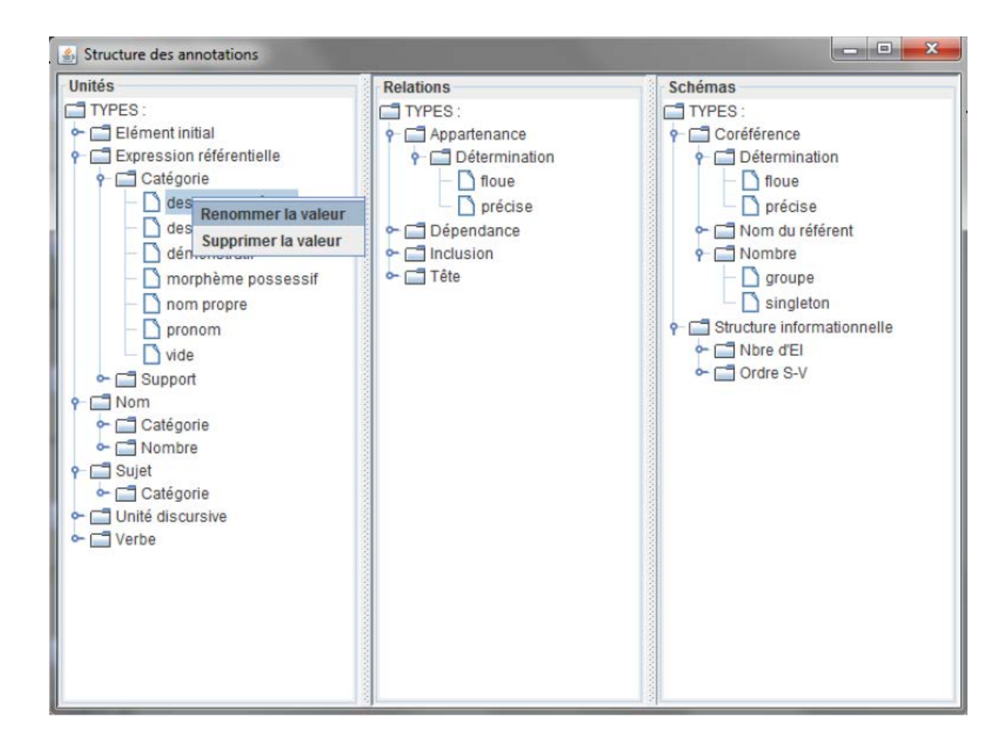
Sous cet intitulé sont regroupées les actions portant sur la structure d'annotation, aussi appelé modèle d'annotation.

- **Gestion de la structure…**

En sélectionnant « gestion de la structure », l'utilisateur accède au schéma d'annotation.

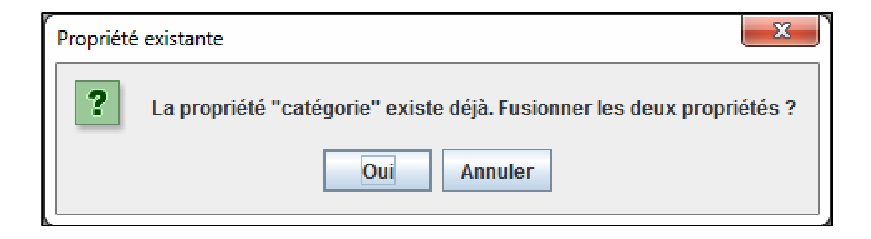
La fenêtre est divisée en 3 parties, qui correspondent aux trois niveaux d'annotation :
- unités
- relations
- schéma

Pour chacun de ces niveaux, l'utilisateur peut *renommer* ou *supprimer* : types, élément et valeur – par simple clic droit sur l'entité souhaitée.



! Si l'entité que l'utilisateur souhaite *supprimer* est présente dans les annotations, un message d'avertissement sera affiché.

! Si le nouveau nom choisi pour une entité est déjà présent dans le schéma d'annotation, un message d'avertissement sera affiché. L'utilisateur pourra alors fusionner s'il le souhaite les 2 entités. Ce type de manipulation peut être très utile lors du travail d'annotation, Analec permettant en effet de fusionner de manière dynamique et ergonomique 2 entités.



- **Ouvrir…**

  Sélectionner un fichier structure existant, déjà créé par un utilisateur Analec : *.ecs

- **Enregistrer…**

Permet d'enregistrer dans un fichier indépendant le schéma d'annotation : *.ecs.

Quelle utilité pour le fichier .ecs ?

- Ce fichier permet à un groupe travaillant sur un même projet, une même problématique d'utiliser le même modèle d'annotation.
- Il contient le nom des champs et les valeurs qui y sont associées. Il peut être ouvert dans éditeur de tableur.

A noter :

Vous ouvrez un fichier déjà annoté, faut-il ou non ouvrir le fichier structure (.ecv) ?

- Si le fichier .ecs porte le même nom que le fichier .ec, il est ouvert automatiquement.
- Si le fichier .ecs a déjà été ouvert avec le fichier .ec, alors la structure du fichier est celle du fichier .ecs ouvert précédemment.
- Si le fichier .ecs n'a jamais été ouvert avec le fichier .ec, alors la structure du fichier ouvert comportera uniquement les *types*, *propriétés* et *valeurs* instanciés dans le fichier .ec. Il faut ouvrir le fichier .ecs souhaité pour avoir à disposition la totalité du modèle d'annotation.

! Une étoile apparaît à la suite du nom du fichier quand les dernières modifications apportées ne sont pas enregistrées.

! Lors d'un premier travail d'annotation ou de toute modification de structure, il est important d'enregistrer la structure d'annotation.

- **Importer un modèle Glozz…**

Permet d'importer dans l'outil un fichier *.aam contenant un modèle d'annotation mis en place sous le logiciel Glozz.

- **Exporter un modèle Glozz…**

Permet d'exporter le modèle d'annotation du fichier courant au format Glozz.
Cette action entraîne la création d'un fichier *.aam.

c) Vue

Utilité du fichier *vue* :

Il permet à chacun de personnaliser son interface, son formulaire d'annotation.

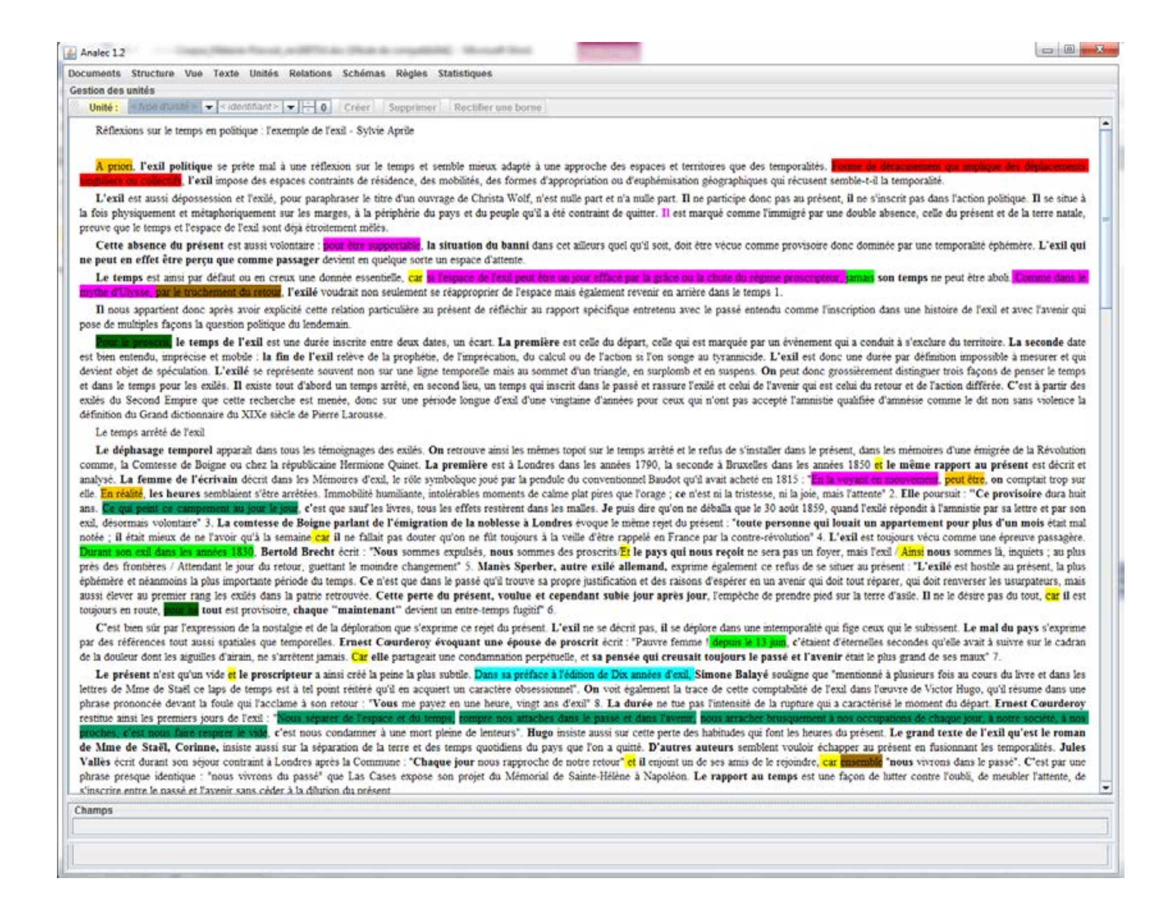
A l'inverse, il permet aussi à un groupe travaillant avec une même structure, d'échanger un environnement de travail, une certaine mise en forme des champs d'annotations et annotations.

Contenu du fichier *vue* :

Il permet la mise en forme des éléments annotés.

Dans la copie d'écran ci-dessous (cf. annotation EIOMSIT) :

- les sujets sont grasseyés.
- les éléments qui précèdent le sujet sont en surbrillance – de couleur différente selon la valeur de l'élément annoté.



- **Gestion de la vue…**

  […]

- **Ouvrir (.ecv)…**

Sélectionner un fichier vue existant, déjà créé par un utilisateur Analec : *.ecv

- **Enregistrer (.ecv)…**
  […]

  ! Si les modifications apportées dans la vue ne sont pas enregistrées, au dessus de la barre d'outils est noté « vue modifiée » suivi d'une étoile :

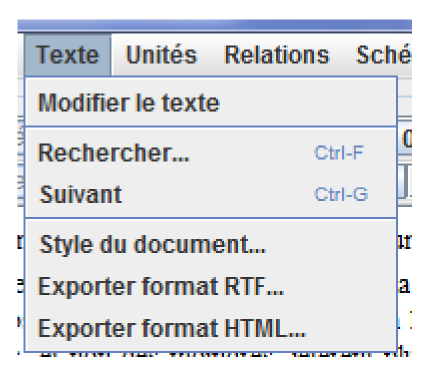


- **Rétablir la vue par défaut**

  [définir vue par défaut]

d) Texte

[3 zones…]



- **Modifier le texte**
  Permet de modifier le texte (ex : correction de faute d'orthographe, insérer des commentaires). Il faut utiliser cette fonctionnalité avec grande précaution car elle modifie le texte, et aucun retour en arrière n'est possible après avoir cliqué sur « valider ».

- **Rechercher…**

! **Le raccourci CTL-F existe**

- **Suivant**

! **Le raccourci CTL-G existe**

- **Style du document…**
  Permet d'agir sur la présentation à l'écran du document :

modifier la taille des caractères, augmenter l'interligne, mettre en gras, etc…

! Mise en forme conservé dans le fichier vue à l'enregistrement du fichier vue.

- **Exporter format RTF…**
  Ce format est pratique pour la relecture. Il permet par exemple de visualiser le texte avec toutes les mises en forme contenues dans le fichier .ecv.

  ! Exporter en .rtf, c'est exporter le fichier tel qu'il tel qu'il apparaît à l'écran.

- **Exporter format HTML…**
  […]

e) Unités

[…]

f) Relations

[…]

g) Règles

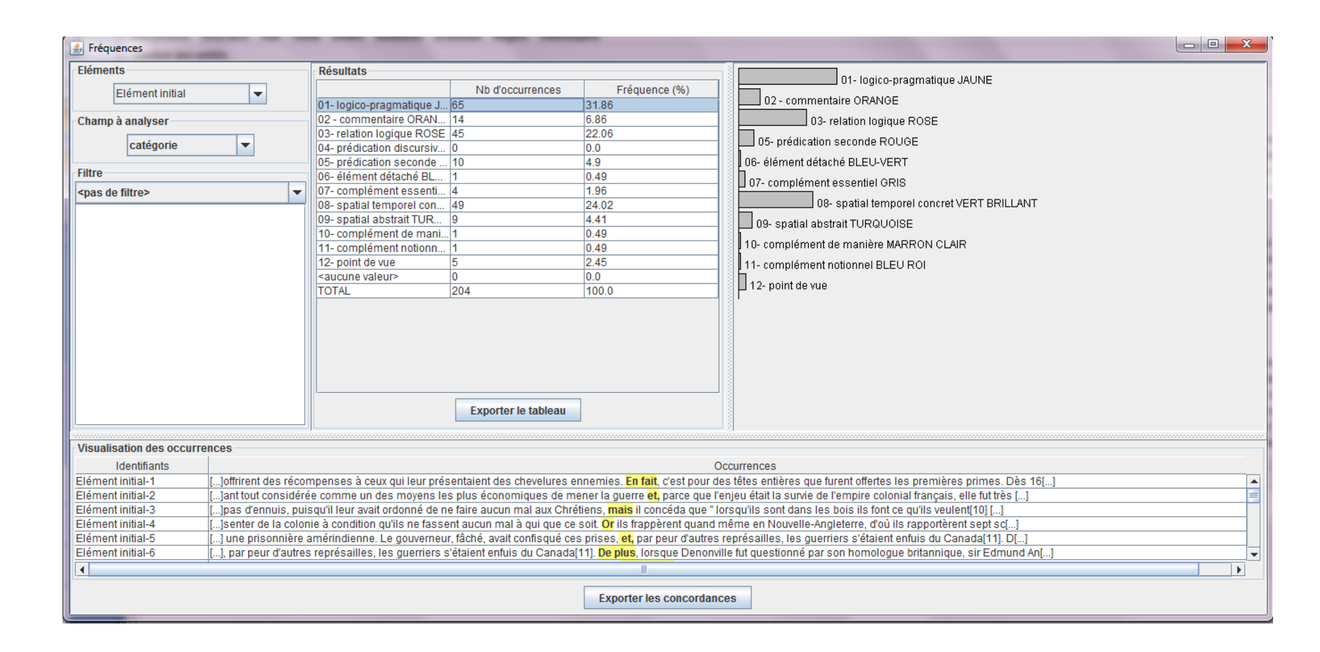
[…]

h) Statistiques

Une analyse quantitative des unités est accessible par l'onglet « statistiques ».
L'ensemble des valeurs des propriétés de chaque unité est accessible et permet à tout moment d'observer les données dans leur ensemble.

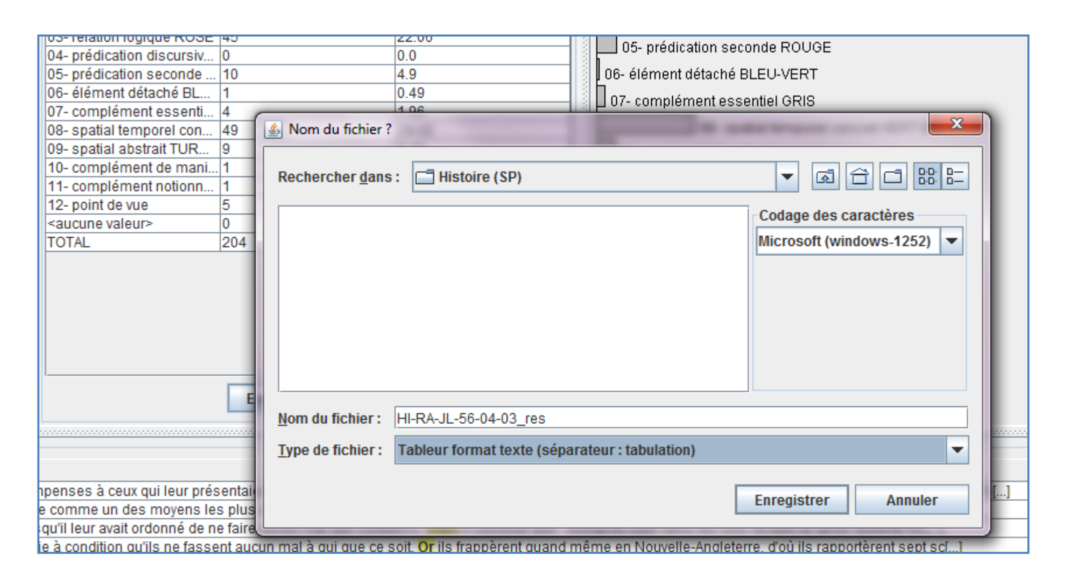
3 types de traitement statistiques sont disponibles :

* **Fréquences…** (calculs de fréquence)



Ce menu permet 2 types d'export - export sous forme de tableaux :

- "**Exporter le tableau**" : le tableau présent au dessus de cet onglet sera exporté au format tableur, séparation des colonnes par tabulation – comme précisé à l'écran.



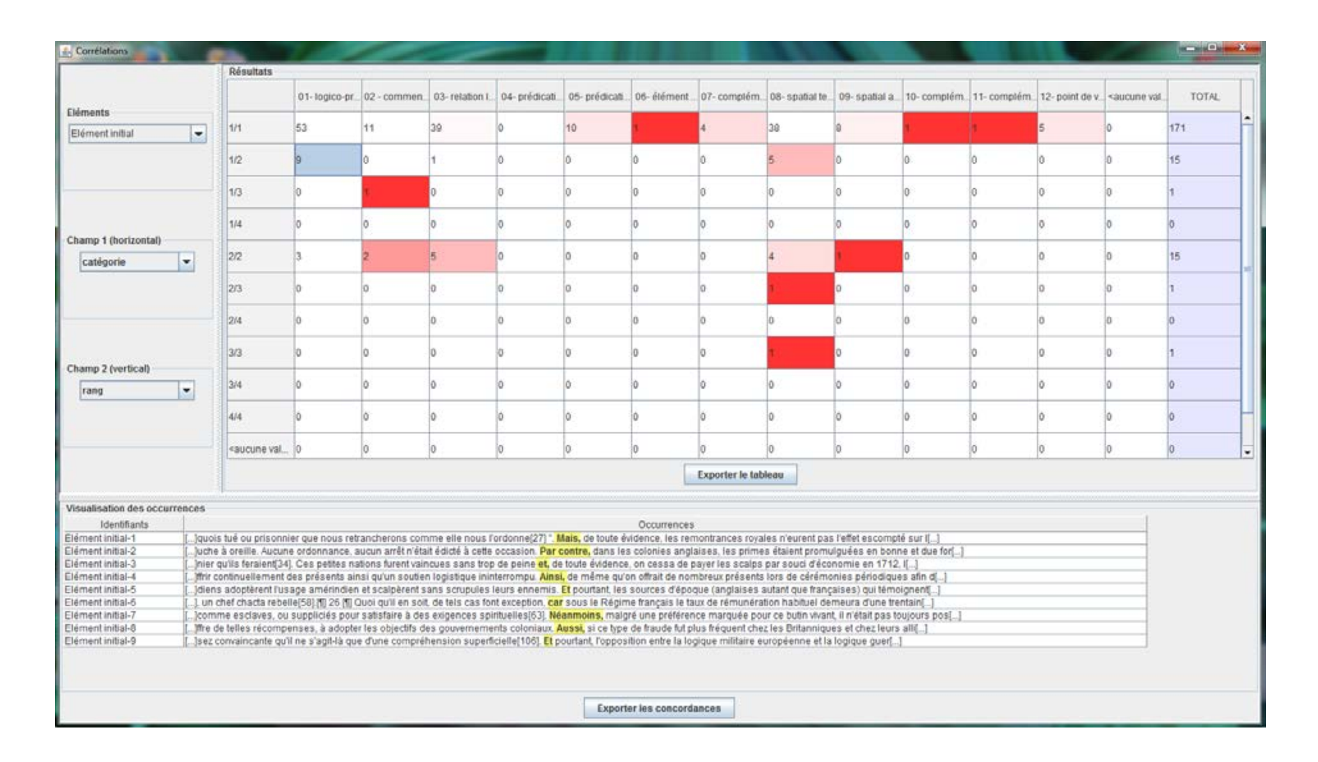
- "**exporter les concordance**s" : export du tableau de concordance

(qui est *actif* uniquement si un type d'annotation est choisi (colonne n°1 dans le tableau résultat)

A Noter :

Ces exports peuvent être ouverts dans des logiciels permettant la lecture de tableurs (ex : Excel, …) et y être manipulés selon les besoins – création de graphes par exemple.

* **Corrélations…** (calculs de corrélations)



	01- logico-pr...	02 - commen...	03- relation l...	04- prédicati...	05- prédicati...	06- élément ...	07- complém...	08- spatial te...	09- spatial a...	10- complém...	11- complém...	12- point de v...	<aucune val...	TOTAL
1/1	53	11	39	0	10	1	4	38	8	1	1	5	0	171
1/2	9	0	1	0	0	0	0	5	0	0	0	0	0	15
1/3	0	1	0	0	0	0	0	0	0	0	0	0	0	1
1/4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2/2	3	2	5	0	0	0	0	4	1	0	0	0	0	15
2/3	0	0	0	0	0	0	0	1	0	0	0	0	0	1
2/4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3/3	0	0	0	0	0	0	0	1	0	0	0	0	0	1
3/4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4/4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<aucune val...	0	0	0	0	0	0	0	0	0	0	0	0	0	0

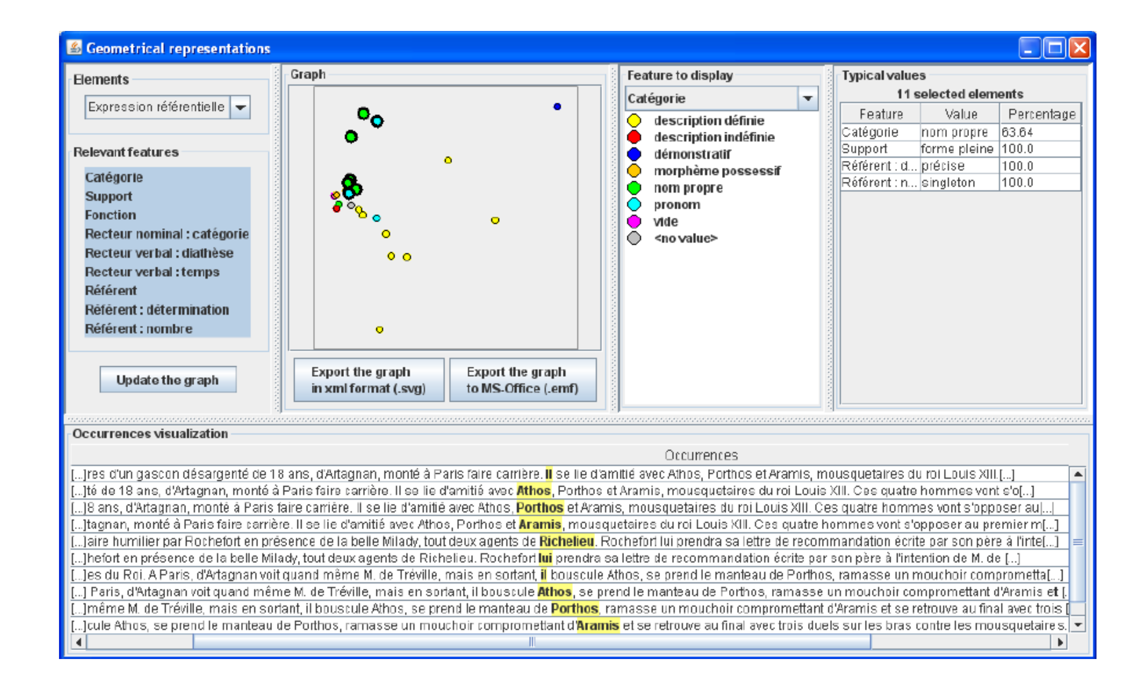
Parmi l'ensemble des valeurs d'un élément (« Eléments ») donné, l'utilisateur choisi 2 valeurs (« champ1 » et « champ2 »).

[…]

Ce menu permet un export de type tableau.

* **Représentations géométriques…** (analyses factorielles de correspondance via des représentations géométriques sous la forme de nuages de points)

Ce menu permet un export de type graphe, au format (svg ou emf).



! Quand un champ est sélectionné (« champ pertinent »), penser à actualiser le graphe en cliquant sur "mettre à jour le graphe".

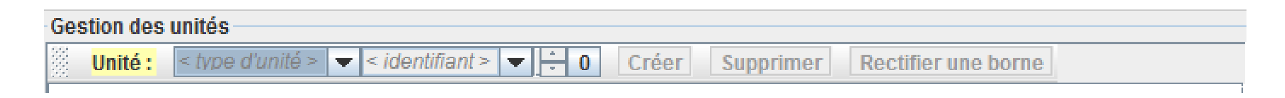
! Pour avoir un tableau de concordance, cliquer sur un des points du graphe.

! Si un champ est sélectionné (menu "champ à afficher"), alors un code couleur est appliqué pour chacune de ses valeurs (arbitraire, non géré par l'utilisateur...).

2) La barre des tâches

[varie selon action en cours]

a) Annoter des unités



Unité

1° sélectionner un type d’unité

2° surligner dans le texte la zone de texte souhaité

3° cliquer sur créer

! Il est possible de passer par l’étape 2° avant l’étape 1°.
! C’est la barre des tâches active par défaut.
! Après diverses manipulations, pour revenir à ce type de barre des tâches :

Unités > Gestion des unités

[<type d’unité>]

[<identifiant>]

Les flèches 

Permet de se déplacer d’une unité annotée à une autre :

- Soit la précédente
- Soit la suivante

Créer

Supprimer

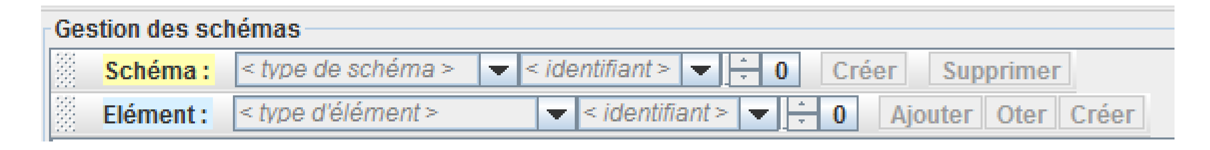
Rectifier une borne

1°sélectionner un élément annoté

2°cliquer sur "rectifier une borne"

3° cliquer sur le nouvel endroit qui sera la nouvelle borne de l'élément initialement délimité différemment

b) Annoter des schémas



[...]

III. Travailler avec Analec

1) Fichier texte brut

procéder dans l'ordre suivant :

1° importer le fichier

2° ouvrir ou importer un modèle (fichier .ecs ou .aam)

3° ouvrir une vue (fichier .ecv)

2) Ouvrir un fichier déjà annoté

Vous avez commencé à annoter un fichier dans une version d'Analec donnée et vous souhaitez continuer à annoter ce fichier dans Analec.

- **Sans changer le modèle d'annotation**

! Si le fichier annoté est dans un dossier où se trouve un fichier .ecv (c'est-à-dire une vue) ayant le même nom que le fichier annoté (à l'extension prêt !), alors cette vue est

ouverte automatiquement :

Exemple :

Vous travaillez sur le fichier : LI-LO-JD-XX-04-04.ec

Votre vue se nomme : LI-LO-JD-XX-04-04.ecv

- Utiliser un nouveau modèle d'annotation

[…]

[message -> Un message apparaît "le modèle va être fusionné avec le modèle existant" > cliquer sur 'oui']

3) Quelques fonctionnalités pratiques

a) Les touches de raccourci

CTL M (*memorize*)

Mémorise les champs et valeurs qui viennent d'être implémentées.
Faire CTL M une fois l'unité annotée permet de mémoriser l'ensemble des valeurs des champs associées à cette unité.

CTL R (restore)

Redistribue les champs et valeur précédemment enregistrés.
Faire CTL R une fois l'unité créée permet de redistribuer à l'unité en question les valeurs précédemment enregistrée.
! Il faut donc que ce soit le même *type* d'unité.

Tab

Dans le formulaire d'annotation, la tabulation permet de passer d'un champ à un autre.

F1 et **F2**

Ces 2 touchent permettent de passer d'une unité annotée à l'autre
F1 : en allant vers le bas
F2 : en allant vers le haut

b) Annoter un ensemble d'exemples à valeur « nulle ».

Lors du travail d'annotation, l'utilisateur peut décider de délimiter une chaîne de caractères en lui attribuant un type, puis une propriété, mais sans pour autant savoir quelle valeur choisir. La chaîne en question prend alors une valeur *nulle*.

Il est possible qu'au fil du travail un ensemble d'exemples soit ainsi annotés, et qu'à un moment du travail l'indécision soit levée. L'utilisateur sait la valeur qu'il veut attribuer à ces exemples.

En allant dans *structure > gestion de la structure*, l'utilisateur peut modifier cette valeur *nulle* par une des valeurs du type. Il lui suffit de choisir la valeur qu'il souhaite attribuer aux chaînes non encore annotées comme valeur par défaut (2 clics droit sur cette valeur).

Alors le message suivant apparaît :



Cliquer sur *oui*. A noter :

- dès que vous cliquez sur oui, les valeurs *nulles* prennent pour valeur la valeur par défaut sélectionnée. Aucun retour en arrière possible.
- Il est donc possible de suite après cette manipulation, de décider de ne pas laisser la valeur en question par défaut pour la suite du travail, de choisir ou non une autre valeur par défaut.

IV. Cas pratiques et exemples

[…]

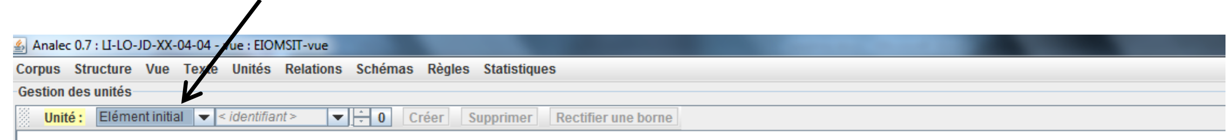
1) EIOMSIT et Analec.

[…] *mettre quelques mots sur EIOMSIT*

a) Visualiser un type d'unité.

Dans le menu initial, sélectionner un type d'unité. Le type d'unité choisi, déjà annoté dans le texte, est surligné en jaune.

Par exemple : "élément initial".



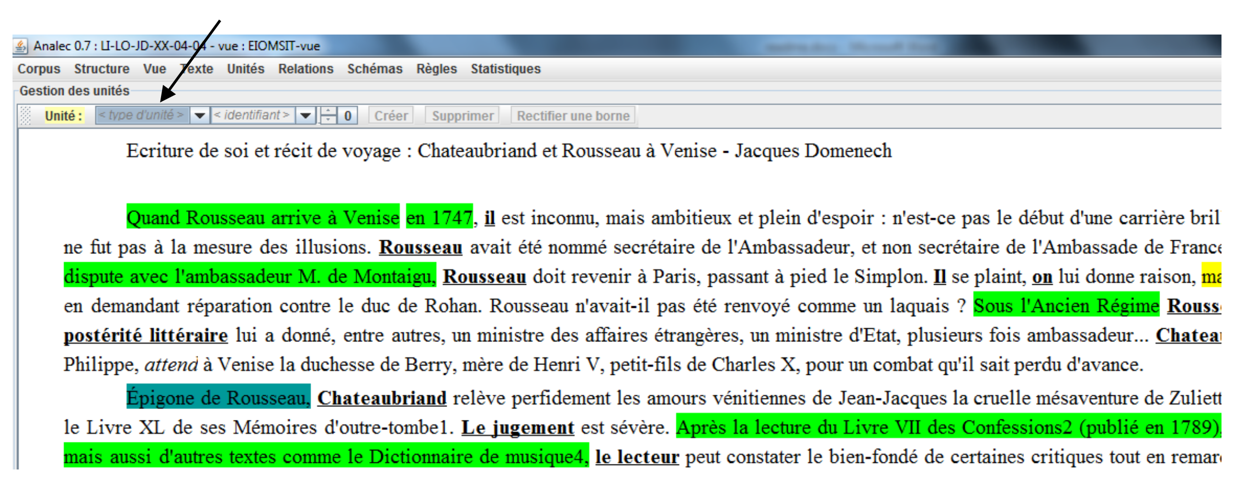
Ecriture de soi et récit de voyage : Chateaubriand et Rousseau à Venise - Jacques Domenech

Quand Rousseau arrive à Venise en 1747, il est inconnu, mais ambitieux et plein d'espoir : n'est-ce pas le début d'une carrière br
fut pas à la mesure des illusions. Rousseau avait été nommé secrétaire de l'Ambassadeur, et non secrétaire de l'Ambassade de France au
avec l'ambassadeur M. de Montaigu, Rousseau doit revenir à Paris, passant à pied le Simplon. Il se plaint, on lui donne raison, mai
demandant réparation contre le duc de Rohan. Rousseau n'avait-il pas été renvoyé comme un laquais ? Sous l'Ancien Régime Rousseau
littéraire lui a donné, entre autres, un ministre des affaires étrangères, un ministre d'Etat, plusieurs fois ambassadeur... Chateaubrianc
attend à Venise la duchesse de Berry, mère de Henri V, petit-fils de Charles X, pour un combat qu'il sait perdu d'avance.

Épigone de Rousseau, Chateaubriand relève perfidement les amours vénitiennes de Jean-Jacques la cruelle mésaventure de Zulie
Livre XL de ses Mémoires d'outre-tombe1. Le jugement est sévère. Après la lecture du Livre VII des Confessions2 (publié en 1789),
aussi d'autres textes comme le Dictionnaire de musique4, le lecteur peut constater le bien-fondé de certaines critiques tout en remarc

b) Visualiser l'ensemble des unités annotées.

Pour visualiser l'ensemble des unités annotées – quand un fichier *vue* lui est associé – il faut n'avoir sélectionné <u>aucun type d'unité</u> dans le bandeau supérieur.



! Un export au format .rtf de ce document, donnera lieu à un texte avec des éléments surlignés.

2) Problème de compatibilité entre 2 formats Analec

(d'une ancienne à une nouvelle version…)

Actuellement, le passage à une nouvelle version Analec ne pose pas de problème de compatibilité avec la version précédente. Cependant, il a pu arriver qu'un travail d'annotation fait dans une ancienne version d'Analec ne puisse être téléchargé dans la nouvelle version de l'outil. Comment faire ?

Il suffit de passer par le format Glozz (mais inutile d'avoir installé Glozz !)

1° ouvrir Analec dans la version antérieure (celle sous laquelle le fichier a été créé) : faire un export sous Glozz du travail fait.

Corpus > Exporter des données Glozz

(*Signifie* « exporter des données sous Glozz »)

Une série de fichiers est générée :

.ac (le corpus)

.aa (les annotations à appliquer sur le fichier)

.aam (le modèle d'annotation)

2° dans la nouvelle version importer les fichiers Glozz :

Pour avoir le texte annoté, importer sous Analec deux fichiers :

Corpus > Importer des données Glozz >

Choisir le corpus souhaité : * .ac

-> Automatiquement il demande le fichier contenant les annotations. Choisir le fichier .aa approprié.

Bibliographie

Quelques articles abordant l'utilisation d'Analec :

- Mélanie-Becquet F., Landragin F., *Linguistique outillée pour l'étude des chaînes de référence : questions méthodologiques et solutions techniques*, Langages, 195, 2014, pp. 117-137.
- Landragin F., Tanguy N., *Référence et coréférence du pronom indéfini on*, Langages 195, Armand Colin, Paris, 2014, pp. 99-115.
- Landragin F., Poibeau T., Victorri B., ANALEC: a New Tool for the Dynamic Annotation of Textual Data Eighth International Conference on Language Resources and Evaluation, Istanbul, Turquey, 2012, pp. 357-362.
- Landragin F., *Une procédure d'analyse et d'annotation des chaînes de coréférence dans des textes écrits*, Corpus 10, http://corpus.revues.org/2010, 2011, pp. 61-80.